

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- AG
- (i) APPLICANT: Gonsalves, Dennis
Pang, Sheng-Zhi
- (ii) TITLE OF INVENTION: TOMATO SPOTTED WILT VIRUS
- (iii) NUMBER OF SEQUENCES: 30
- (iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Nixon Peabody LLP
(B) STREET: Clinton Square, P.O. Box 1051
(C) CITY: Rochester
(D) STATE: New York
(E) COUNTRY: U.S.A.
(F) ZIP: 14603
- (v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: US 08/495,484
(B) FILING DATE: 27-JAN-1994
(C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Goldman, Michael L.
(B) REGISTRATION NUMBER: 30,727
(C) REFERENCE/DOCKET NUMBER: 19603/10303
- (ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (716) 263-1304
(B) TELEFAX: (716) 263-1600

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

AGCAGGCAAA ACTCGCAGAA CTTGC

(2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 25 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

GCAAGTTCTG CGAGTTTTC CTGCT

25

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 32 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

AGCTAACCAT GGTAAAGCTC ACTAAGGAAA GC

32

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 32 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

AGCATTCCAT GGTTAACACA CTAAGCAAGC AC

32

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 2216 base pairs
 - (B) TYPE: nucleic acid

662207 "E0292160

(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

| | | | | | | |
|-------------|------------|------------|-------------|------------|------------|------|
| CAAGTTGAAA | GCAACAACAG | AACTGTAAAT | TCTCTTGCAG | TGAAATCTCT | GCTCATGTCA | 60 |
| GCAGAAAACA | ACATCATGCC | TAActCTCAA | GCTTCCACTG | ATTCTCATT | CAAGCTGAGC | 120 |
| CTCTGGCTAA | GGGTTCCAAA | GGTTTTGAAG | CAGGTTTCCA | TTCAGAAATT | GTTCAAGGTT | 180 |
| GCAGGAGATG | AAACAAACAA | AACATTTTAT | TTATCTATTG | CCTGCATTCC | AAACCATAAC | 240 |
| AGTGTTGAGA | CAGCTTTAAA | CATTACTGTT | ATTTGCAAGC | ATCAGCTCCC | AATTCGCAAA | 300 |
| TGCAAAGCTC | CTTTTGAATT | ATCAATGATG | TTTTCTGATT | TAAAGGAGCC | TTACAACATT | 360 |
| GTTTCATGACC | CTTCATACCC | CAAAGGATCG | GTTCCAATGC | TCTGGCTCGA | AACTCACACA | 420 |
| TCTTTGCACA | AGTTCTTTGC | AACTAACTTG | CAAGAAGATG | TAATCATCTA | CACTTTGAAC | 480 |
| AACCTTGAGC | TAActCCTGG | AAAGTTAGAT | TTAGGTGAAA | GAACCTTGAA | TTACAGTGAA | 540 |
| GATGCCTACA | AAAGGAAATA | TTTCCTTTCA | AAAACACTTG | AATGTCTTCC | ATCTAACACA | 600 |
| CAAActATGT | CTTACTTAGA | CAGCATCCAA | ATCCCTTCAT | GGAAGATAGA | CTTTGCCAGA | 660 |
| GGAGAAATTA | AAATTTCTCC | ACAATCTATT | TCAGTTGCAA | AATCTTTGTT | AAAGCTTGAT | 720 |
| TTAAGCGGGA | TCAAAAAGAA | AGAATCTAAG | GTAAAGGAAG | CGTATGCTTC | AGGATCAAAA | 780 |
| TAATCTTGCT | TTGTCCAGCT | TTTTCTAATT | ATGTTATGTT | TATTTTCTTT | CTTTACTTAT | 840 |
| AATTATTTCT | CTGTTTGTCA | TCTCTTTCAA | ATTCCTCCTG | TCTAGTAGAA | ACCATAAAAA | 900 |
| CAAAAAATAA | AAATGAAAAT | AAAATTAAAA | TAAAATAAAA | TCAAAAAATG | AAATAAAAAC | 960 |
| AACAAAAAAT | TAAAAAACGA | AAAACCAAAA | AGACCCGAAA | GGGACCAATT | TGGCCAAATT | 1020 |
| TGGGTTTTGT | TTTTGTTTTT | TGTTTTTTGT | TTTTTATTTT | TTATTTTATT | TTTATTTTAT | 1080 |
| TTTATTTTTA | TTTTATTTTT | ATTTTATTTA | TTTTTTGTTT | TCGTTGTTTT | TGTTATTTTA | 1140 |
| TTATTTATTA | AGCACAACAC | ACAGAAAGCA | AACTTTAATT | AAACACACTT | ATTTAAAATT | 1200 |
| TAACACACTA | AGCAAGCACA | AGCAATAAAG | ATAAAGAAAAG | CTTTATATAT | TTATAGGCTT | 1260 |
| TTTTATAAAT | TAActTACAG | CTGCTTTCAA | GCAAGTTCTG | CGAGTTTTCG | CTGCTTTTTA | 1320 |
| ACCCCGAACA | TTTCATAGAA | CTTGTTAAGA | GTTTCACTGT | AATGTTCCAT | AGCAACACTC | 1380 |
| CCTTTAGCAT | TAGGATTGCT | GGAGCTAAGT | ATAGCAGCAT | ACTCTTTCCT | CTTCTTCACC | 1440 |

00136703-1033000

| | |
|--|------|
| TGATCTTCAT TCATTTCAAA TGCTTTGCTT TTCAGCACAG TGCAAACTTT TCCTAAGGCT | 1500 |
| TCCTTGGTGT CATACTTCTT TGGGTCGATC CCGAGGTCCT TGTATTTTGC ATCCTGATAT | 1560 |
| ATAGCCAAGA CAACACTGAT CATCTCAAAG CTATCAACTG AAGCAATAAG AGGTAAGCTA | 1620 |
| CCTCCCAGCA TTATGGCAAG TCTCACAGAC TTTGCATCAT CGAGAGGTAA TCCATAGGCT | 1680 |
| TGAATCAAAG GATGGGAAGC AATCTTAGAT TTGATAGTAT TGAGATTCTC AGAATTCCCA | 1740 |
| GTTTCTTCAA CAAGCCTGAC CCTGATCAAG CTATCAAGCC TTCTGAAGGT CATGTCAGTG | 1800 |
| CCTCCAATCC TGTCTGAAGT TTTCTTTATG GTAATTTTAC CAAAAGTAAA ATCGCTTTGC | 1860 |
| TTAATAACCT TCATTATGCT CTGACGATTC TTTAGGAATG TCAGACATGA AATAACGCTC | 1920 |
| ATCTTCTTGA TCTGGTCGAT GTTTTCCAGA CAAAAGTCTT TGAAGTTGAA TGCTACCAGA | 1980 |
| TTCTGATCTT CCTCAAACCTC AAGGTCTTTG CCTTGTGTCA ACAAAGCAAC AATGCTTTCC | 2040 |
| TTAGTGAGCT TAACCTTAGA CATGATGATC GTAAAAGTTG TTATATGCTT TGACCGTATG | 2100 |
| TAACCTCAAGG TGCGAAAGTG CAACTCTGTA TCCCGCAGTC GTTCTTAGG TTCTTAATGT | 2160 |
| GATGATTTGT AAGACTGAGT GTTAAGGTAT GAACACAAAA TTGACACGAT TGCTCT | 2216 |

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1709 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

| | |
|---|-----|
| AAATTCTCTT GCAGTGAAAT CTCTGCTCAT GTTAGCAGAA AACAAATCA TGCCTAACTC | 60 |
| TCAAGCTTTT GTCAAAGCTT CTACTGATTC TAATTTCAAG CTGAGCCTCT GGCTAAGGGT | 120 |
| TCCAAAGGTT TTGAAGCAGA TTTCCATTCA GAAATTGTTC AAGGTTGCAG GAGATGAAAC | 180 |
| AAATAAAACA TTTTATTTAT CTATTGCCTG CATTCCAAAC CATAACAGTG TTGAGACAGC | 240 |
| TTTAAACATT ACTGTTATTT GCAAGCATCA GCTCCCAATT CGTAAATGTA AAACCTCTTT | 300 |
| TGAATTATCA ATGATGTTTT CTGATTTAAA GGAGCCTTAC AACATTATTC ATGATCCTTC | 360 |
| ATATCCCCAA AGGATTGTTC ATGCTCTGCT TGAAACTCAC ACATCTTTTG CACAAGTTCT | 420 |

| | |
|--|------|
| TTGCAACAAC TTGCAAGAAG ATGTGATCAT CTACACCTTG AACCAACCATG AGCTAACTCC | 480 |
| TGGAAAGTTA GATTTAGGTG AAATAACTTT GAATTACAAT GAAGACGCCT ACAAAGGAA | 540 |
| ATATTTCTTT TCAAAAACAC TTGAATGTCT TCCATCTAAC ATACAAACTA TGTCTTATTT | 600 |
| AGACAGCATC CAAATCCCTT CCTGGAAGAT AGACTTTGCC AGGGGAGAAA TAAAATTTT | 660 |
| TCCACAATCT ATTTCAAGTTG CAAAATCTTT GTTAAATCTT GATTTAAGCG GGATTAAAA | 720 |
| GAAAGAATCT AAGATTAAGG AAGCATATGC TTCAGGATCA AAATGATCTT GCTGTGTCCA | 780 |
| GCTTTTTCTA ATTATGTTAT GTTTATTTTC TTTCTTTACT TATAATTATT TTTCTGTTTG | 840 |
| TCATTTCTTT CAAATTCCTC CTGTCTAGTA GAAACCATAA AAACAAAAAT AAAAATAAAA | 900 |
| TAAATCAAA ATAAAATAAA AATCAAAAAA TGAAATAAAA GCAACAAAAA AATTAAAAAA | 960 |
| CAAAAAACCA AAAAAGATCC CGAAAGGACA ATTTTGGCCA AATTTGGGGT TTGTTTTTGT | 1020 |
| TTTTTGTTTT TTTGTTTTTT GTTTTTATTT TTATTTTTAT TTTTATTTTT ATTTTATTTT | 1080 |
| ATTTTATGTT TTTGTTGTTT TTGTTATTTT GTTATTTATT AAGCACAACA CACAGAAAGC | 1140 |
| AACTTTAAT TAAACACACT TATTTAAAT TTAACACACT AAGCAAGCAC AAACAATAAA | 1200 |
| GATAAGAAA GCTTTATATA TTTATAGGCT TTTTATAAAT TTAACCTACA GCTGCTTTTA | 1260 |
| AGCAAGTTCT GTGAGTTTTG CCTGTTTTTT AACCCCAAAC ATTTCATAGA ACTTGTTAAG | 1320 |
| GGTTTCACTG TAATGTTCCA TAGCAATACT TCCTTTAGCA TTAGGATTGC TGGAGCTAAG | 1380 |
| TATAGCAGCA TACTCTTTC CCTTCTTCAC CTGATCTTCA TTCATTTCAA ATGCTTTTCT | 1440 |
| TTTCAGCACA GTGCAAACTT TTCCTAAGGC TTCCCTGGTG TCATACTTCT TTGGGTCGAT | 1500 |
| CCCGAGATCC TTGTATTTTG CATCCTGATA TATAGCCAAG ACAACACTGA TCATCTCAAA | 1560 |
| GCTATCAACT GAAGCAATAA GAGGTAAGCT ACCTCCCAGC ATTATGGCAA GCCTCACAGA | 1620 |
| CTTTGCATCA TCAAGAGGTA ATCCATAGGC TTGAATCAAA GGGTGGGAAG CAATCTTAGA | 1680 |
| TTTGATAGTA TTGAGATTCT CAGAATTCC | 1709 |

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 260 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted May 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

| | | | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Gln | Val | Glu | Ser | Asn | Asn | Arg | Thr | Val | Asn | Ser | Leu | Ala | Val | Lys | Ser | 1 | 5 | 10 | 15 |
| Leu | Leu | Met | Ser | Ala | Glu | Asn | Asn | Ile | Met | Pro | Asn | Ser | Gln | Ala | Ser | 20 | 25 | 30 | |
| Thr | Asp | Ser | His | Phe | Lys | Leu | Ser | Leu | Trp | Leu | Arg | Val | Pro | Lys | Val | 35 | 40 | 45 | |
| Leu | Lys | Gln | Val | Ser | Ile | Gln | Lys | Leu | Phe | Lys | Val | Ala | Gly | Asp | Glu | 50 | 55 | 60 | |
| Thr | Asn | Lys | Thr | Phe | Tyr | Leu | Ser | Ile | Ala | Cys | Ile | Pro | Asn | His | Asn | 65 | 70 | 75 | 80 |
| Ser | Val | Glu | Thr | Ala | Leu | Asn | Ile | Thr | Val | Ile | Cys | Lys | His | Gln | Leu | 85 | 90 | 95 | |
| Pro | Ile | Arg | Lys | Cys | Lys | Ala | Pro | Phe | Glu | Leu | Ser | Met | Met | Phe | Ser | 100 | 105 | 110 | |
| Asp | Leu | Lys | Glu | Pro | Tyr | Asn | Ile | Val | His | Asp | Pro | Ser | Tyr | Pro | Lys | 115 | 120 | 125 | |
| Gly | Ser | Val | Pro | Met | Leu | Trp | Leu | Glu | Thr | His | Thr | Ser | Leu | His | Lys | 130 | 135 | 140 | |
| Phe | Phe | Ala | Thr | Asn | Leu | Gln | Glu | Asp | Val | Ile | Ile | Tyr | Thr | Leu | Asn | 145 | 150 | 155 | 160 |
| Asn | Leu | Glu | Leu | Thr | Pro | Gly | Lys | Leu | Asp | Leu | Gly | Glu | Arg | Thr | Leu | 165 | 170 | 175 | |
| Asn | Tyr | Ser | Glu | Asp | Ala | Tyr | Lys | Arg | Lys | Tyr | Phe | Leu | Ser | Lys | Thr | 180 | 185 | 190 | |
| Leu | Glu | Cys | Leu | Pro | Ser | Asn | Thr | Gln | Thr | Met | Ser | Tyr | Leu | Asp | Ser | 195 | 200 | 205 | |
| Ile | Gln | Ile | Pro | Ser | Trp | Lys | Ile | Asp | Phe | Ala | Arg | Gly | Glu | Ile | Lys | 210 | 215 | 220 | |
| Ile | Ser | Pro | Gln | Ser | Ile | Ser | Val | Ala | Lys | Ser | Leu | Leu | Lys | Leu | Asp | 225 | 230 | 235 | 240 |
| Leu | Ser | Gly | Ile | Lys | Lys | Lys | Glu | Ser | Lys | Val | Lys | Glu | Ala | Tyr | Ala | 245 | 250 | 255 | |
| Ser | Gly | Ser | Lys | 260 | | | | | | | | | | | | | | | |

662207 662206

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 858 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

TTAACACACT AAGCAAGCAC AAACAATAAA GATAAAGAAA GCTTTATATA TTTATAGGCT 60
TTTTTATAAT TTAAC TTACA GCTGCTTTTA AGCAAGTTCT GTGAGTTTTG CCTGTTTTTT 120
AACCCCAAAC ATTT CATAGA ACTTGTTAAG GGTTTCACTG TAATGTTCCA TAGCAATACT 180
TCCTTTAGCA TTAGGATTGC TGGAGCTAAG TATAGCAGCA TACTCTTTCC CCTTCTTCAC 240
CTGATCTTCA TTCATTTCAA ATGCTTTTCT TTTCAGCACA GTGCAAAC TT TCCCTAAGGC 300
TTCCCTGGTG TCATACTTCT TTGGGTCGAT CCCGAGATCC TTGTATTTTG CATCCTGATA 360
TATAGCCAAG ACAACACTGA TCATCTCAAA GCTATCAACT GAAGCAATAA GAGGTAAGCT 420
ACCTCCAGC ATTATGGCAA GCCTCACAGA CTTTGCATCA TCAAGAGGTA ATCCATAGGC 480
TTGACTCAAA GGGTGGGAAG CAATCTTAGA TTTGATAGTA TTGAGATTCT CAGAATTCCC 540
AGTTTCCTCA ACAAGCCTGA CCCTGATCAA GCTATCAAGC CTTCTGAAGG TCATGTCAGT 600
GGCTCCAATC CTGTCTGAAG TTTTCTTTAT GGTAATTTTA CCAAAAGTAA AATCGCTTTG 660
CTTAATAACC TTCATTATGC TCTGACGATT CTTCAGGAAT GTCAGACATG AAATAATGCT 720
CATCTTTTTG ATCTGGTCAA GGTTTTCCAG ACAAAAAGTC TTGAAGTTGA ATGCTACCAG 780
ATTCTGATCT TCCTCAAAC CAAGGTCTTT GCCTTGTGTC AACAAAGCAA CAATGCTTTC 840
CTTAGTGAGC TTAACCAT 858

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2028 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

| | |
|--|------|
| AAATTCTCTT GCAGTGAAAT CTCTGCTCAT GTTAGCAGAA AACAAATCA TGCCTAACTC | 60 |
| TCAAGCTTTT GTCAAAGCTT CTACTGATTC TAATTTCAAG CTGAGCCTCT GGCTAAGGGT | 120 |
| TCCAAAGGTT TTGAAGCAGA TTTCCATTCA GAAATTGTTC AAGGTTGCAG GAGATGAAAC | 180 |
| AAATAAAACA TTTTATTTAT CTATTGCCTG CATTCCAAAC CATAACAGTG TTGAGACAGC | 240 |
| TTTAAACATT ACTGTTATTT GCAAGCATCA GCTCCCAATT CGTAAATGTA AAATCCTTTT | 300 |
| TGAATTATCA ATGATGTTTT CTGATTTAAA GGAGCCTTAC AACATTATTC ATGATCCTTC | 360 |
| ATATCCCCAA AGGATTGTTC ATGCTCTGCT TGAAACTCAC ACATCTTTTG CACAAGTTCT | 420 |
| TTGCAACAAC TTGCAAGAAG ATGTGATCAT CTACACCTTG AACCAACCATG AGCTAACTCC | 480 |
| TGGAAAGTTA GATTTAGGTG AAATAACTTT GAATTACAAT GAAGACGCCT ACAAAGGAA | 540 |
| ATATTTCTTT TCAAAAACAC TTGAATGTCT TCCATCTAAC ATACAAACTA TGTCTTATTT | 600 |
| AGACAGCATC CAAATCCCTT CCTGGAAGAT AGACTTTGCC AGGGGAGAAA TTAAAATTTT | 660 |
| TCCACAATCT ATTTAGTTG CAAAATCTTT GTTAAATCTT GATTTAAGCG GGATTAAAAA | 720 |
| GAAAGAATCT AAGATTAAGG AAGCATATGC TTCAGGATCA AAATGATCTT GCTGTGTCCA | 780 |
| GCTTTTTCTA ATTATGTTAT GTTTATTTTC TTTCTTTACT TATAATTATT TTTCTGTTTG | 840 |
| TCATTTCTTT CAAATTCCTC CTGTCTAGTA GAAACCATAA AAACAAAAAT AAAAATAAAA | 900 |
| TAAATCAAA ATAAAATAAA AATCAAAAAA TGAAATAAAA GCAACAAAAA AATTAAAAAA | 960 |
| CAAAAAACCA AAAAAGATCC CGAAAGGACA ATTTTGGCCA AATTTGGGGT TTGTTTTTGT | 1020 |
| TTTTTGTTTT TTTGTTTTTT GTTTTTATTT TTATTTTAT TTTTATTTT ATTTTATTTT | 1080 |
| ATTTTATGTT TTTGTTGTTT TTGTTATTTT GTTATTTATT AAGCACAACA CACAGAAAGC | 1140 |
| AACTTTAAT TAAACACACT TATTTAAAT TTAACACACT AAGCAAGCAC AAACAATAAA | 1200 |
| GATAAAGAAA GCTTTATATA TTTATAGGCT TTTTATAAT TTAACCTACA GCTGCTTTTA | 1260 |
| AGCAAGTTCT GTGAGTTTTG CCTGTTTTTT AACCCCAAAC ATTCATAGA ACTTGTTAAG | 1320 |
| GGTTTCACTG TAATGTTCCA TAGCAATACT TCCTTTAGCA TTAGGATTGC TGGAGCTAAG | 1380 |
| TATAGCAGCA TACTCTTCC CCTTCTTCAC CTGATCTTCA TTCATTTCAA ATGCTTTTCT | 1440 |
| TTTCAGCACA GTGCAAACTT TTCCTAAGGC TTCCCTGGTG TCATACTTCT TTGGGTCGAT | 1500 |
| CCCGAGATCC TTGTATTTTG CATCCTGATA TATAGCCAAG ACAACACTGA TCATCTCAAA | 1560 |
| GCTATCAACT GAAGCAATAA GAGGTAAGCT ACCTCCCAGC ATTATGGCAA GCCTCACAGA | 1620 |

662207-00230160

CTTTGCATCA TCAAGAGGTA ATCCATAGGC TTGACTCAAA GGGTGGGAAG CAATCTTAGA 1680
TTTGATAGTA TTGAGATTCT CAGAATTCCC AGTTTCCTCA ACAAGCCTGA CCCTGATCAA 1740
GCTATCAAGC CTTCTGAAGG TCATGTCAGT GGCTCCAATC CTGTCTGAAG TTTTCTTTAT 1800
GGTAATTTTA CCAAAAGTAA AATCGCTTTG CTTAATAACC TTCATTATGC TCTGACGATT 1860
CTTCAGGAAT GTCAGACATG AAATAATGCT CATCTTTTTG ATCTGGTCAA GGTTTTCCAG 1920
ACAAAAAGTC TTGAAGTTGA ATGCTACCAG ATTCTGATCT TCCTCAAACCT CAAGGTCTTT 1980
GCCTTGTC AACAAAGCAA CAATGCTTTC CTTAGTGAGC TTAACCAT 2028

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

TTCTGGTCTT CTTCAAACCTC A 21

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

CTGTAGCCAT GAGCAAAG 18

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 467 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

662207 03232400

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

```

Met Ser Ser Gly Val Tyr Glu Ser Ile Ile Gln Thr Lys Ala Ser Val
1          5          10          15

Trp Gly Ser Thr Ala Ser Gly Lys Ser Ile Val Asp Ser Tyr Trp Ile
          20          25          30

Tyr Glu Phe Pro Thr Gly Ser Pro Leu Val Gln Thr Gln Leu Tyr Ser
          35          40          45

Asp Ser Arg Ser Lys Ser Ser Phe Gly Tyr Thr Ser Lys Ile Gly Asp
50          55          60

Ile Pro Ala Val Glu Glu Glu Ile Leu Ser Gln Asn Val His Ile Pro
65          70          75          80

Val Phe Asp Asp Ile Asp Phe Ser Ile Asn Ile Asn Asp Ser Phe Leu
          85          90          95

Ala Ile Ser Val Cys Ser Asn Thr Val Asn Thr Asn Gly Val Lys His
          100          105          110

Gln Gly His Leu Lys Val Leu Ser Leu Ala Gln Leu His Pro Phe Glu
          115          120          125

Pro Val Met Ser Arg Ser Glu Ile Ala Ser Arg Phe Arg Leu Gln Glu
          130          135          140

Glu Asp Ile Ile Pro Asp Asp Lys Tyr Ile Ser Ala Ala Asn Lys Gly
145          150          155          160

Ser Leu Ser Cys Val Lys Glu His Thr Tyr Lys Val Glu Met Ser His
          165          170          175

Asn Gln Ala Leu Gly Lys Val Asn Val Leu Ser Pro Asn Arg Asn Val
          180          185          190

His Glu Trp Leu Tyr Ser Phe Lys Pro Asn Phe Asn Gln Ile Glu Ser
          195          200          205

Asn Asn Arg Thr Val Asn Ser Leu Ala Val Lys Ser Leu Leu Met Ala
          210          215          220

Thr Glu Asn Asn Ile Met Pro Asn Ser Gln Ala Phe Val Lys Ala Ser
225          230          235          240

Thr Asp Ser His Phe Lys Leu Ser Leu Trp Leu Arg Ile Pro Lys Val
          245          250          255

Leu Lys Gln Ile Ala Ile Gln Lys Leu Phe Lys Phe Ala Gly Asp Glu

```

662207" 66230160

| | | |
|--|-----|-----|
| 260 | 265 | 270 |
| Thr Gly Lys Ser Phe Tyr Leu Ser Ile Ala Cys Ile Pro Asn His Asn | | |
| 275 | 280 | 285 |
| Ser Val Glu Thr Ala Leu Asn Val Thr Val Ile Cys Arg His Gln Leu | | |
| 290 | 295 | 300 |
| Pro Ile Pro Lys Ser Lys Ala Pro Phe Glu Leu Ser Met Ile Phe Ser | | |
| 305 | 310 | 315 |
| Asp ² Leu Lys Glu Pro Tyr Asn Thr Val His Asp Pro Ser Tyr Pro Gln | | |
| 325 | 330 | 335 |
| Arg Ile Val His Ala Leu Leu Glu Thr His Thr Ser Phe Ala Gln Val | | |
| 340 | 345 | 350 |
| Leu Cys Asn Lys Leu Gln Glu Asp Val Ile Ile Tyr Thr Ile Asn Ser | | |
| 355 | 360 | 365 |
| Pro Glu Leu Thr Pro Ala Lys Leu Asp Leu Gly Glu Arg Thr Leu Asn | | |
| 370 | 375 | 380 |
| Tyr Ser Glu Asp Ala Ser Lys Lys Lys Tyr Phe Leu Ser Lys Thr Leu | | |
| 385 | 390 | 395 |
| Glu Cys Leu Pro Val Asn Val Gln Thr Met Ser Tyr Leu Asp Ser Ile | | |
| 405 | 410 | 415 |
| Gln Ile Pro Ser Trp Lys Ile Asp Phe Ala Arg Gly Glu Ile Arg Ile | | |
| 420 | 425 | 430 |
| Ser Pro Gln Ser Thr Pro Ile Ala Arg Ser Leu Leu Lys Leu Asp Leu | | |
| 435 | 440 | 445 |
| Ser Lys Ile Lys Glu Lys Lys Ser Leu Thr Trp Glu Thr Ser Ser Tyr | | |
| 450 | 455 | 460 |
| Asp Leu Glu | | |
| 465 | | |

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 258 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

| | | | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Ser | Lys | Val | Lys | Leu | Thr | Lys | Glu | Asn | Ile | Val | Ser | Leu | Leu | Thr | 1 | 5 | 10 | 15 |
| Gln | Ser | Ala | Asp | Val | Glu | Phe | Glu | Glu | Asp | Gln | Asn | Gln | Val | Ala | Phe | 20 | 25 | 30 | |
| Asn | Phe | Lys | Thr | Phe | Cys | Gln | Glu | Asn | Leu | Asp | Leu | Ile | Lys | Lys | Met | 35 | 40 | 45 | |
| Ser | Ile | Thr | Ser | Cys | Leu | Thr | Phe | Leu | Lys | Asn | Arg | Gln | Gly | Ile | Met | 50 | 55 | 60 | |
| Lys | Val | Val | Asn | Gln | Ser | Asp | Phe | Thr | Phe | Gly | Lys | Val | Thr | Ile | Lys | 65 | 70 | 75 | 80 |
| Lys | Asn | Ser | Glu | Arg | Val | Gly | Ala | Lys | Asp | Met | Thr | Phe | Arg | Arg | Leu | 85 | 90 | 95 | |
| Asp | Ser | Met | Ile | Arg | Val | Lys | Leu | Ile | Glu | Glu | Thr | Ala | Asn | Asn | Glu | 100 | 105 | 110 | |
| Asn | Leu | Ala | Ile | Ile | Lys | Ala | Lys | Ile | Ala | Ser | His | Pro | Leu | Val | Gln | 115 | 120 | 125 | |
| Ala | Tyr | Gly | Leu | Pro | Leu | Ala | Asp | Ala | Lys | Ser | Val | Arg | Leu | Ala | Ile | 130 | 135 | 140 | |
| Met | Leu | Gly | Gly | Ser | Ile | Pro | Leu | Ile | Ala | Ser | Val | Asp | Ser | Phe | Glu | 145 | 150 | 155 | 160 |
| Met | Ile | Ser | Val | Val | Leu | Ala | Ile | Tyr | Gln | Asp | Ala | Lys | Tyr | Lys | Glu | 165 | 170 | 175 | |
| Leu | Gly | Ile | Glu | Pro | Thr | Lys | Tyr | Asn | Thr | Lys | Glu | Ala | Leu | Gly | Lys | 180 | 185 | 190 | |
| Val | Cys | Thr | Val | Leu | Lys | Ser | Lys | Gly | Phe | Thr | Met | Asp | Asp | Ala | Gln | 195 | 200 | 205 | |
| Ile | Asn | Lys | Gly | Lys | Glu | Tyr | Ala | Lys | Ile | Leu | Ser | Ser | Cys | Asn | Pro | 210 | 215 | 220 | |
| Asn | Ala | Lys | Gly | Ser | Ile | Ala | Met | Asp | Tyr | Tyr | Ser | Asp | Asn | Leu | Asp | 225 | 230 | 235 | 240 |
| Lys | Phe | Tyr | Glu | Met | Phe | Gly | Val | Lys | Lys | Glu | Ala | Lys | Ile | Ala | Gly | 245 | 250 | 255 | |
| Val | Ala | | | | | | | | | | | | | | | | | | |

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 3049 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

| | |
|--|------|
| AGAGCAATTG GGTCAATTTT TATTCTAAAT CGAACCTCAA CTAGCAAATC TCAGAACTGT | 60 |
| AATAAGCACA AGAGCACAAG AGCCACAATG TCATCAGGTG TTTATGAATC GATCATTGAG | 120 |
| ACAAAGGCTT CAGTTTGGGG ATCGACAGCA TCTGGTAAGT CCATCGTGGA TTCTTACTGG | 180 |
| ATTTATGAGT TTCCAACCTGG TTCTCCACTG GTTCAAACCTC AGTTGTACTC TGATTGAGG | 240 |
| AGCAAAAGTA GCTTCGGCTA CACTTCAAAA ATTGGTGATA TTCCTGCTGT AGAGGAGGAA | 300 |
| ATTTTATCTC AGAACGTTCA TATCCCAGTG TTTGATGATA TTGATTTTCA CATCAATATC | 360 |
| AATGATTCTT TCTTGGCAAT TTCTGTTTGT TCCAACACAG TTAACACCAA TGGAGTGAAG | 420 |
| CATCAGGGTC ATCTTAAAGT TCTTTCTCTT GCCCAATTGC ATCCCTTTGA ACCTGTGATG | 480 |
| AGCAGGTCAG AGATTGCTAG CAGATTCCGG CTCCAAGAAG AAGATATAAT TCCTGATGAC | 540 |
| AAATATATAT CTGCTGCTAA CAAGGGATCT CTCTCCTGTG TCAAAGAACA TACTTACAAA | 600 |
| GTCGAAATGA GCCACAATCA GGCTTTAGGC AAAGTGAATG TTCTTTCTCC TAACAGAAAT | 660 |
| GTTTCATGAGT GGCTGTATAG TTTCAAACCA AATTTCAACC AGATCGAAAG TAATAACAGA | 720 |
| ACTGTAAATF CTCTTGCACT CAAATCTTTG CTCATGGCTA CAGAAAACAA CATTATGCCT | 780 |
| AACTCTCAAG CTTTTGTAA AGCTTCTACT GATTCTCATT TTAAGTTGAG CCTTTGGCTG | 840 |
| AGAATTCCAA AAGTTTTGAA GCAAATAGCC ATACAGAAGC TCTTCAAGTT TGCAGGAGAC | 900 |
| GAAACCGGTA AAAGTTTCTA TTTGTCTATT GCATGCATCC CAAATCACAA CAGTGTGGAA | 960 |
| ACAGCTTTAA ATGTCACCTG TATATGTAGA CATCAGCTTC CAATCCCTAA GTCCAAAGCT | 1020 |
| CCTTTTGAAT TATCAATGAT TTTCTCCGAT CTGAAAGAGC CTTACAACAC TGTGCATGAT | 1080 |
| CCTTCATATC CTCAAAGGAT TGTTTCATGCT TTGCTTGAGA CTCACACTTC CTTTGCACAA | 1140 |
| GTTCTCTGCA ACAAGCTGCA AGAAGATGTG ATCATATATA CTATAAACAG CCCTGAACTA | 1200 |
| ACCCAGCTA AGCTGGATCT AGGTGAAAGA ACCTTGAACCT ACAGTGAAGA TGCTTCGAAG | 1260 |
| AAGAAGTATT TTCTTTCAAA AACACTCGAA TGCTTGCCAG TAAATGTGCA GACTATGTCT | 1320 |
| TATTTGGATA GCATCCAGAT TCCTTCATGG AAGATAGACT TTGCCAGAGG AGAGATCAGA | 1380 |

| | | | | | | |
|-------------|------------|------------|------------|------------|-------------|------|
| ATCTCCCCCTC | AATCTACTCC | TATTGCAAGA | TCTTTGCTCA | AGCTGGATTT | GAGCAAGATC | 1440 |
| AAGGAAAAGA | AGTCCTTGAC | TTGGGAAACA | TCCAGCTATG | ATCTAGAATA | AAAGTGGCTC | 1500 |
| ATACTACTCT | AAGTAGTATT | TGTCAACTTG | CTTATCCTTT | ATGTTGTTTA | TTTCTTTTAA | 1560 |
| ATCTAAAGTA | AGTTAGATTC | AAGTAGTTTA | GTATGCTATA | GCATTATTAC | AAAAAATACA | 1620 |
| AAAAAATACA | AAAAAATACA | AAAAATATAA | AAAACCCAAA | AAGATCCCAA | AAGGGACGAT | 1680 |
| TTGGTTGATT | TACTCTGTTT | TAGGCTTATC | TAAGCTGCTT | TTGTTTGAGC | AAAATAACAT | 1740 |
| TGTAACATGC | AATAACTGGA | ATTTAAAGTC | CTAAAAGAAG | TTTCAAAGGA | CAGCTTAGCC | 1800 |
| AAAATTGGTT | TTTGTTTTTG | TTTTTTTGTT | TTTTGTTTTT | TTGTTTTATT | TTTATTTTTA | 1860 |
| GTTTATTTTT | TGTTTTTGTT | ATTTTTATTT | TTATTTTATT | TTCTTTTATT | TTATTTATAT | 1920 |
| ATATATCAAA | CACAATCCAC | ACAAATAATT | TTAATTTCAA | ACATTCTACT | GATTTAACAC | 1980 |
| ACTTAGCCTG | ACTTTATCAC | ACTTAACACG | CTTAGTTAGG | CTTTAACACA | CTGAAC TGAA | 2040 |
| TTAAACACAC | CTTAGTATTA | TGCATCTCTT | AATTAACACA | CTTTAATAAT | ATGCATCTCT | 2100 |
| GAATCAGCCT | TAAAGAAGCT | TTTATGCAAC | ACCAGCAATC | TTGGCCTCTT | TCTTAACTCC | 2160 |
| AAACATTTCA | TAGAATTTGT | CAAGATTATC | ACTGTAATAG | TCCATAGCAA | TGCTTCCCTT | 2220 |
| AGCATTGGGA | TTGCAAGAAC | TAAGTATCTT | GGCATATTCT | TTCCCTTTGT | TTATCTGTGC | 2280 |
| ATCATCCATT | GTAAATCCTT | TGCTTTTAAG | CACTGTGCAA | ACCTTCCCCA | GAGCTTCCTT | 2340 |
| AGTGTTGTAC | TTAGTTGGTT | CAATCCCTAA | CTCCTTGTA | TTTGCACTCT | GATATATGGC | 2400 |
| AAGAACAACA | CTGATCATCT | CGAAGCTGTC | AACAGAAGCA | ATGAGAGGGA | TACTACCTCC | 2460 |
| AAGCATTATA | GCAAGTCTCA | CAGATTTTGC | ATCTGCCAGA | GGCAGCCCGT | AAGCTTGGAC | 2520 |
| CAAAGGGTGG | GAGGCAATTT | TTGCTTTGAT | AATAGCAAGA | TTCTCATTGT | TTGCAGTCTC | 2580 |
| TTCTATGAGC | TTCACTCTTA | TCATGCTATC | AAGCCTCCTG | AAAGTCATAT | CCTTAGCTCC | 2640 |
| AACTCTTTCA | GAATTTTTCT | TTATCGTGAC | CTTACCAAAA | GTAAAATCAC | TTTGGTTCAC | 2700 |
| AACTTTCATA | ATGCCTTGGC | GATTCTTCAA | GAAAGTCAAA | CATGAAGTGA | TACTCATTTT | 2760 |
| CTTAATCAGG | TCAAGATTTT | CCTGACAGAA | AGTCTTAAAG | TTGAATGCGA | CCTGGTTCTG | 2820 |
| GTCTTCTTCA | AACTCAACAT | CTGCAGATTG | AGTTAAAAGA | GAGACAATGT | TTTCTTTTGT | 2880 |
| GAGCTTGACC | TTAGACATGG | TGGCAGTTTA | GATCTAGACC | TTTCTCGAGA | GATAAGATTC | 2940 |
| AAGGTGAGAA | AGTGCAACAC | TGTAGACCGC | GGTCGTTACT | TATCCTGTTA | ATGTGATGAT | 3000 |

004667 1090

TTGTATTGCT GAGTATTAGG TTTTGAATA AAATTGACAC AATTGCTCT

3049

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 777 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

| | |
|--|-----|
| TTATGCAACA CCAGCAATCT TGGCCTCTTT CTTAACTCCA AACATTTTCAT AGAATTTGTC | 60 |
| AAGATTATCA CTGTAATAGT CCATAGCAAT GCTTCCCTTA GCATTGGGAT TGCAAGAACT | 120 |
| AAGTATCTTG GCATATTCTT TCCCTTTGTT TATCTGTGCA TCATCCATTG TAAATCCTTT | 180 |
| GCTTTTAAGC ACTGTGCAAA CCTTCCCCAG AGCTTCCTTA GTGTTGTACT TAGTTGGTTC | 240 |
| AATCCCTAAC TCCTTGTA CTGTCATCTTG ATATATGGCA AGAACAACAC TGATCATCTC | 300 |
| GAAGCTGTCA ACAGAAGCAA TGAGAGGGAT ACTACCTCCA AGCATTATAG CAAGTCTCAC | 360 |
| AGATTTTGCA TCTGCCAGAG GCAGCCCGTA AGCTTGGACC AAAGGGTGGG AGGCAATTTT | 420 |
| TGCTTTGATA ATAGCAAGAT TCTCATTGTT TGCAGTCTCT TCTATGAGCT TCACTCTTAT | 480 |
| CATGCTATCA AGCCTCCTGA AAGTCATATC CTTAGCTCCA ACTCTTTCAG AATTTTCTT | 540 |
| TATCGTGACC TTACCAAAG TAAATCACT TTGGTTCACA ACTTTCATAA TGCCTTGGCG | 600 |
| ATTCTTCAAG AAAGTCAAAC ATGAAGTGAT ACTCATTTTC TTAATCAGGT CAAGATTTTC | 660 |
| CTGACAGAAA GTCTTAAAGT TGAATGCGAC CTGGTTCTGG TCTTCTTCAA ACTCAACATC | 720 |
| TGCAGATTGA GTTAAAAGAG AGACAATGTT TTCTTTTGTT AGCTTGACCT TAGACAT | 777 |

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

GTTCTGAGAT TTGCTAGT

18

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

TTATATCTTC TTCTTGGA

18

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1401 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

| | |
|--|-----|
| ATGTCATCAG GTGTTTATGA ATCGATCATT CAGACAAAGG CTTCAGTTTG GGGATCGACA | 60 |
| GCATCTGGTA AGTCCATCGT GGATTCTTAC TGGATTTATG AGTTTCCAAC TGGTTCTCCA | 120 |
| CTGGTTCAAA CTCAGTTGTA CTCTGATTCTG AGGAGCAAAA GTAGCTTCGG CTACACTTCA | 180 |
| AAAATTGGTG ATATTCCTGC TGTAGAGGAG GAAATTTTAT CTCAGAACGT TCATATCCCA | 240 |
| GTGTTTGATG ATATTGATTT CAGCATCAAT ATCAATGATT CTTTCTTGGC AATTTCTGTT | 300 |
| TGTTCCAACA CAGTTAACAC CAATGGAGTG AAGCATCAGG GTCATCTTAA AGTTCTTTCT | 360 |
| CTTGCCCAAT TGCATCCCTT TGAACCTGTG ATGAGCAGGT CAGAGATTGC TAGCAGATTC | 420 |
| CGGCTCCAAG AAGAAGATAT AATTCCTGAT GACAAATATA TATCTGCTGC TAACAAGGGA | 480 |
| TCTCTCTCCT GTGTCAAAGA ACATACTTAC AAAGTCGAAA TGAGCCACAA TCAGGCTTTA | 540 |
| GGCAAAGTGA ATGTTCTTTC TCCTAACAGA AATGTTTCATG AGTGGCTGTA TAGTTTCAAA | 600 |
| CCAAATTTCA ACCAGATCGA AAGTAATAAC AGAACTGTAA ATTCTCTTGC AGTCAAATCT | 660 |

| | |
|---|------|
| TTGCTCATGG CTACAGAAAA CAACATTATG CCTAACTCTC AAGCTTTTGT TAAAGCTTCT | 720 |
| ACTGATTCTC ATTTTAAGTT GAGCCTTTGG CTGAGAATTC CAAAAGTTTT GAAGCAAATA | 780 |
| GCCATACAGA AGCTCTTCAA GTTTGCAGGA GACGAAACCG GTAAAAGTTT CTATTTGTCT | 840 |
| ATTGCATGCA TCCCAAATCA CAACAGTGTG GAAACAGCTT TAAATGTCAC TGTTATATGT | 900 |
| AGACATCAGC TTCCAATCCC TAAGTCCAAA GCTCCTTTTG AATTATCAAT GATTTTCTCC | 960 |
| GATCTGAAAG AGCCTTACAA CACTGTGCAT GATCCTTCAT ATCCTCAAAG GATTGTTCAT | 1020 |
| GCTTTGCTTG AGACTCACAC TTCCTTTGCA CAAGTTCTCT GCAACAAGCT GCAAGAAGAT | 1080 |
| GTGATCATAT ATACTATAAA CAGCCCTGAA CTAACCCAG CTAAGCTGGA TCTAGGTGAA | 1140 |
| AGAACCTTGA ACTACAGTGA AGATGCTTCG AAGAAGAAGT ATTTTCTTTC AAAAACACTC | 1200 |
| GAATGCTTGC CAGTAAATGT GCAGACTATG TCTTATTTGG ATAGCATCCA GATTCCTTCA | 1260 |
| TGGAAGATAG ACTTTGCCAG AGGAGAGATC AGAATCTCCC CTCAATCTAC TCCTATTGCA | 1320 |
| AGATCTTTGC TCAAGCTGGA TTTGAGCAAG ATCAAGGAAA AGAAGTCCTT GACTTGGGAA | 1380 |
| ACATCCAGCT ATGATCTAGA A | 1401 |

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 777 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

| | |
|---|-----|
| ATGTCTAAGG TCAAGCTCAC AAAAGAAAAC ATTGTCTCTC TTTTAACTCA ATCTGCAGAT | 60 |
| GTTGAGTTTG AAGAAGACCA GAACCAGGTC GCATTCAACT TTAAGACTTT CTGTCAGGAA | 120 |
| AATCTTGACC TGATTAAGAA AATGAGTATC ACTTCATGTT TGACTTTCTT GAAGAATCGC | 180 |
| CAAGGCATTA TGAAAGTTGT GAACCAAAGT GATTTTACTT TTGGTAAGGT CACGATAAAG | 240 |
| AAAAATTCTG AAAGAGTTGG AGCTAAGGAT ATGACTTTCA GGAGGCTTGA TAGCATGATA | 300 |
| AGAGTGAAGC TCATAGAAGA GACTGCAAAC AATGAGAATC TTGCTATTAT CAAAGCAAAA | 360 |
| ATTGCCTCCC ACCCTTTGGT CCAAGCTTAC GGGCTGCCTC TGGCAGATGC AAAATCTGTG | 420 |
| AGACTTGCTA TAATGCTTGG AGGTAGTATC CCTCTCATTG CTTCTGTTGA CAGCTTCGAG | 480 |

ATGATCAGTG TTGTTCTTGC CATATATCAA GATGCAAAGT ACAAGGAGTT AGGGATTGAA 540
CCAACTAAGT ACAACACTAA GGAAGCTCTG GGGAAGGTTT GCACAGTGCT TAAAAGCAAA 600
GGATTTACAA TGGATGATGC ACAGATAAAC AAAGGGAAAG AATATGCCAA GATACTTAGT 660
TCTTGCAATC CCAATGCTAA GGGAAGCATT GCTATGGACT ATTACAGTGA TAATCTTGAC 720
AAATTCTATG AAATGTTTGG AGTTAAGAAA GAGGCCAAGA TTGCTGGTGT TGCATAA 777

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 40 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

TACTTATCTA GAACCATGGA CAAAGCAAAG ATTACCAAGG 40

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 42 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

TACAGTGGAT CCATGGTTAT TTCAAATAAT TTATAAAAGC AC 42

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 36 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

662204 "E0292460

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

AGCATTGGAT CCATGGTTAA CACACTAAGC AAGCAC

36

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 46 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

AGCTAATCTA GAACCATGGA TGACTCACTA AGGAAAGCAT TGTTC

46

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

CCCCTATCC TTCGCAAGAC CC

22

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 39 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

TACAGTGGAT CCATGGTTAA GGTAATCCAT AGGCTTGAC

39

(2) INFORMATION FOR SEQ ID NO:26:

662207" 662207" 662207"

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 40 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

AGCTAACCAT GGTTAAGCTC ACTAAGGAAA GCATTGTTGC

40

(2) INFORMATION FOR SEQ ID NO:27:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 46 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

AGCTAATCTA GAACCATGGA TGACTCACTA AGGAAAGCAT TGTTGC

46

(2) INFORMATION FOR SEQ ID NO:28:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 36 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

AGCATTGGAT CCATGGTTAA CACACTAAGC AAGCAC

36

(2) INFORMATION FOR SEQ ID NO:29:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 39 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

00125783-1022090

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

TACAGTTCTA GAACCATGGA TGATGCAAAG TCTGTGAGG

39

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 49 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

AGATTCTCTA GACCATGGTG ACTTGATGAG CAAAGTCTGT GAGGCTTGC

49

00128703-1032000

As
Cont